

RAE/REF Metrics Panel

Alan Bundy
UKCRC REF Working Group
RAE 2008 Sub-panel F23 member

RAE 2008 Criteria

- **Environment (20%)**
 - RA5a and standard analyses.
 - Light touch for panels & universities.
- **Esteem (10%)**
 - Indicators: 2a+b. max 4 per CatA.
 - Light touch for panels & universities.
- **Outputs (70%)**
 - 4 outputs per non-ECR CatA, inc textboxes.
 - Heavy touch for panels & universities.
 - But most reliable criterion.

Issues with Metrics

- **Inaccuracy of metrics, e.g.,**
 - lack of coverage by harvesters: iceberg effect,
 - RAE 2008 CS&I panel reported poor correlation,
 - citations & incubation depend on subfield.
- **Potential for behaviour distortion, e.g.,**
 - discourage workshop paper with PhD,
 - focus only on academic impact.
- **Need to apply judgement.**
 - Panels would welcome metric provision,
 - but need to examine outputs.

Assessing Computing Outputs

- **Typical publication pattern:**
 - Initial idea in conference/workshop,
 - High originality, low rigour & significance.
 - Archival journal paper,
 - High rigour, low originality.
 - Application: spin-out, downloads, patent.
 - High significance, low originality.
- **Whole programme needed to reflect 4* quality.**

Assessing Impact

- REF expected to include specific impact measure,
 - e.g., to replace esteem.
- No one metric fully measures:
 - spin-outs, patents, downloads, licences, ...
 - 90% tech transfer via students.
- Need to include *research* impact.
 - Lead times often in *decades*,
 - And then via third party researchers.
- Need to assess *process*,
 - e.g., via narrative.

Need for CS&I Panel

- **Limit to reduction in workload.**
 - Need to exercise *judgement*.
 - Therefore, *examine* outputs, impact, etc.
 - F23 examined ~1200 outputs *each*.
- **Variation in methodology:**
 - maths, engineering, biology, sociology, psychology, philosophy, linguistics, ...
 - No cognate subject reflects this breadth.

Conclusion

- **REF will evolve from RAE 2008,**
 - due to need for *judgement*,
 - rather than *formulaic* use of metrics.
- **Try to reduce output assessment workload,**
 - e.g., fewer per CatA, maybe focus on very best.
 - Must CatA status be tied to output provision?
 - Could assessing *programmes* help?
- **And inclusion of *impact* measures,**
 - which must include *research* impact,
 - and measure of *process*.
- **Requires fine granularity panels,**
 - including CS&I panel.